Secure File Storage on Cloud using Enhanced Hybrid Cryptography

Reece B. D'Souza[1], Dr. Ruby D.[2]

[1]Student, School of Computer Science & Engineering, Vellore Institute of Technology, Tamil Nadu, India
[2]Assistant Professor, School of Computer Science & Engineering, Vellore Institute of Technology, Tamil Nadu, India

---***---

Abstract - *In recent years 'cloud' has taken over the world, offering a new and dynamic form for computation and storage. Several corporations offer cloud services, for example, Amazon Web Services (AWS), Google Cloud Platform (GCP), Microsoft Azure, etc. These cloud services are cheaper in the long run and their services are highly efficient. But a lot of small-scale corporations do not switch to the cloud due to the lack of security associated with the cloud. Therefore, the paper presents a method of storing files in the cloud securely. With the use of modified hybrid encryption, the files stored in the cloud are secure. Accessing any files leaked from the cloud would be futile without the keys and real-time authentication service offered by the model presented in the paper. The model uses the AES algorithm enhanced by threads and RSA Signing and Verification algorithm, along with a real-time OTP generation service to provide high security. While encrypting the file using AES, a key is entered and the file is split into eight pieces and encrypted simultaneously stored in a zip folder on the cloud. While requesting to download a file, the necessary key needs to be entered and an OTP will be emailed to the registered email which needs to be entered. Decryption works in the exact reverse of encryption. Thus, the model offers safe and secure encryption of files and storage in the cloud.*

Key Words: CSPs (Cloud Service Providers), Encryption, Decryption, AES (Advanced Encryption Standard), RSA (Rivest–Shamir–Adleman), OTP (One Time Passwords).

1. INTRODUCTION

Cloud computing is a subscription-based or pay-per-use service that extends an IT company's capability, especially when used in real-time. With the introduction of cloud computing in the early 2000s, there was a significant change in not only the software but other fields as well. Hardware and software are generally used for heavy computation, but now with cloud computing, this load is lifted as the cloud handles it instead.

Cloud computing can be defined as a type of computing that shares computing resources over a network rather than having local or personal servers to handle computation. In a more technical aspect, it is the use of remote servers to store, manage, and process data rather than achieve the same on a local server or a personal computer [1]. Cloud computing offers a variety of different services – such as servers, storage, and computations – which are delivered to the client over the internet. Cloud applications are developed in Platforms as a Service following the PaaS architecture imposed by several providers [1].

Several companies in the world offer these cloud services. They are called 'Cloud Service Providers' or 'CSPs' in short. These companies offer network services, business applications, or even infrastructure in the cloud. The large benefit of using these services via these companies is it is highly efficient when used economically. Rather than maintaining the individual infrastructure of the company, outsourcing these services to third parties for a reasonable price is highly efficient.

The focus of this paper is one of these services, namely storage and more importantly its security. With the introduction of cloud computing, the use of traditional storage devices is declining. This is because, with the help of the cloud, information is stored in a centralized environment accessible from any part of the world having a viable internet connection.

Cloud storage is useful because it eliminates the need for carrying physical storage devices and provides safe backups, while physical devices can be corrupted in some form, leading to loss of the entire data. Even though cloud storage has a higher form of safety over physical devices, it is not completely invulnerable. There have been several cases of data leakage from the cloud through the means of hacking or breaking into the cloud to steal valuable confidential company data for the benefit of the hacker. Even with the availability of data in the cloud, exposing them to applications that may already have security loopholes in them is extremely dangerous [2].

Therefore, storing information on the cloud will simply not suffice. Most businesses have avoided the use of CSPs for the very same reason. There are several methods for securing rough data sets such as the Genetic Algorithm (GA), K Mean Algorithm, KNN (K-Nearest Neighbour) Algorithm, and even Native Bayesian Technique [3]. But since data can be stored in any format in the cloud the use of cryptography would be beneficial. Through the use of cryptography, data can be safely stored. Even though the files are leaked, the files cannot be used as they are encrypted. Cryptography adds another layer to the security of the cloud.

Earlier cryptography was only used for military or diplomatic communication up until the development of public-key cryptography. Cryptography can be defined as a system that through the use of some mathematical algorithm, transforms a sequence of characters that are fed to it [4]. This system is based on the value of a secret key, which is a parameter in the encryption as well as the decryption algorithm. The security of the algorithm depends on this secret key.

There are two forms of cryptography, one being private key cryptography or symmetric cryptography and the other being public key cryptography or asymmetric cryptography. Private key cryptography is a simple form of cryptography where a key is used to encode a message and can only be decoded with that same key. In public-key cryptography, there are two keys namely public and private key. This is often used for two-way communication and in blockchains. Say user A having the public key of user B encodes the message with it and sends it to B. Only the private key of user B can decode this message.

Hybrid cryptography is a combination of both types of cryptography. While the private key algorithm encrypts the message, the other is used for key exchange [5][6]. Hybrid cryptography provides a better form of security than each of two types of cryptosystems as it utilizes the best of both.

The paper presents a model that uses the combination of a private key algorithm called AES and a public key algorithm call RSA algorithm. AES stands for Advanced Encryption Standard. Two Belgian cryptographers, namely Joan Daemen ad Vincent Rijmen developed it in the late 90s, and has become public in the early 2000s [4]. RSA stands for Rivest–Shamir–Adleman, the developers of this algorithm. The encryption key is public and distinct here while its counterpart the private key is kept secret. Due to the slower performance of the AES algorithm, the paper introduces the use of threads to achieve faster performance. It also uses OTP or One Time Passwords to achieve security.

2. LITERATURE SURVEY

A literature survey was conducted on several topics including implementation of cloud services, data security in the cloud, and reviews on several existing models with the same title. Each paper was studied and a brief abstract of the same is written below.

The paper 'Design and Implementation of Cloud Services by using Python' by Priyanka Hariom Singh report about the use of cloud computing using Python Programming Language [1]. Introducing cloud computing and various cloud models, such as Service and Deployment Models, several services by the cloud was introduced. With the help of Docker and Azure, applications are created and run to interact with SQL. The paper gave an introduction to Python, a programming language, and a useful tool for cloud services. An architecture plan was proposed based on the python script. A disadvantage to this setup is that it can prove to be expensive.

The paper 'Data Security in Cloud Computing' by Ahmed Albugmi, Madini O. Alassafi, Robert Walters, and Gary Wills is a review paper for the techniques available in Data Security for data at different stages (i.e. Data at rest, Data in transit, etc.) [2]. It begins assessing the risks and security concerns in cloud computing at various subdomains such as virtualization and storage in the public cloud. It identifies the major security challenges and proposes the means of encryptions via Block Ciphers, Stream Ciphers, and Hash Functions. A drawback to this is the use of all these algorithms and functions could lead to poor performance concerning time.

The paper 'Cloud Computing Security Issues and Its Challenges: A Comprehensive Research' by Jaydip Kumar, reports that the process of cloud storage ensures privacy and security of cloud storage, application interface layer, and accessibility to the cloud user [3]. The cloud user needs not to control or manage the cloud infrastructure including storage, operating system, services, network, and application. The paper further mentions methods for securing datasets using various algorithms such as the Genetic Algorithm, K-Mean Algorithm, KNN Algorithm, and Naïve Bayesian Algorithm. It further presents techniques for securing the cloud such as validation of OTP or One Time Password, Integrity Checking using Provable Data Possession, Access Control, Secure Deletion, Encryption, Data Masking, and Intrusion Detection System.

The paper 'Hybrid Encryption Algorithm' by Gaurav R. Patel and Krunal Panchal, is a review paper for the basic hybrid algorithm [4]. Using an asymmetric algorithm and Diffie Helman Key Exchange, the above mentioned is implemented. It is implemented over a .NET Framework with C# as the programming language. The disadvantage of this is that the Proposed model takes more time than the existing methods.

In the paper 'A Research Paper on New Hybrid Cryptography Algorithm' by Swapnil Chaudhari, Mangesh Pahade, Sahil Bhat, Chetan Jadhav, Tejaswini Sawant, proposes a new Hybrid algorithm, by introducing a substitution method for encryption and decryption at the end and beginning respectively [5]. This method is highly secure, uneasy to decrypt, and highly efficient. This model can be further enhanced with the use of threads or any form of parallelization.

The paper 'A Hybrid Cryptography Technique for Improving Network Security' by V. Kapoor and Rahul Yadav implements Hybrid Cryptography using Java Programming Language [6]. The proposed system provides the same using RSA, DES, and SHA1. It has a better mode of encryption time than the traditional RSA Algorithm. Although the proposed system is excellent in performance, the system does not give freedom for personal passwords.

The paper 'Secure File Storage on Cloud using cryptography' by Joseph Selvanayagam, Akash Singh, Joans Michael, and Jaya Jeswani introduces a framework of hybrid cryptography that combines the use of Data Encryption Standard (DES) and RC-2 (RC2) and storing the encrypted file on the cloud [7]. This is to avoid the threat models that the paper considers being Data Availability, Data Integrity, and Security. With the introduction of the Public Key Cryptosystem, it would allow only the required person to open the file through the means of two-way communication providing a better means of security.

The paper 'Review of Secure File Storage on Cloud using Hybrid Cryptography' by Shruti Kanatt, Amey Jadhav, and Prachi Talwar is a review paper for the algorithms used in Hybrid Cryptography [8]. The paper follows through several existing papers, noting their advantages and their limitations namely AES or Advanced Encryption Standard and Blowfish Algorithm. It provides high security, scalability, and confidentiality. A disadvantage to this is as there is no public key authentication, it may be subject to attack.

The paper 'An Analysis and Survey of Security Techniques for Cloud Computing' by Manish Kumar, Shivani Chauhan, and Ajay Singh is related to the detection of a zombie attack in cloud computing [9]. In the zombie attack, the clones of the virtual machine are created to interact with the cloud server. The various security enhancement techniques are reviewed in this paper in terms of certain parameters. This technique will be improved and then a future proposal can be made for the isolation of zombie attacks in cloud computing.

The paper 'Secure File Storage on Cloud using Hybrid Cryptography' by Aditya Poduval, Abhijeet Doke, Hitesh Nemade, and Rohan Nikam presents a hybrid cryptosystem that involves generating a random key and embedding it into an image using Image Steganography [10]. This same key is used to encrypt a file using the hybrid algorithm. The file after being split into three is encrypted using three different encryption algorithms, namely AES, 3DES, and RC6. Although, the use of Steganography will not always work, as this varies per image the client uses as his profile picture. The performance of the algorithm reduces as time complexity increases. Multithreading would help solve this issue.

3. PROPOSED SYSTEM

The paper presents a new model that implements hybrid cryptography and OTP validation. It was developed using Python, HTML, and Flask. It utilizes two major algorithms, the AES or Advanced Encryption Standard algorithm and the RSA (Rivest–Shamir–Adleman) algorithm. The AES Algorithm is enhanced by using threads. The RSA signing and verification methodology are utilized in this model.

3.1 Algorithms of Components Used

AES Algorithm is based on the 'substitution-permutation network'. It involves several steps of replacing specific inputs with its predefined outputs (substitution) and even shuffling the bits around in any direction (permutation) [11].

In AES, the number of rounds depends on the key length. AES uses 10 rounds for 128-bit keys, 12 rounds for 192-bit, and 14 rounds for 256-bit keys. Each round in encryption involves a sequence of four steps: Byte Substitution (SubBytes), Shifting of Rows, Mixing of Columns, and Adding the Round Key. The decryption process works exactly in reverse. The security of the system will be maintained as long as it is correctly implemented and employs a good key management strategy.

RSA Digital Signatures Algorithm stands apart from the traditional RSA approach. The traditional RSA approach implies encrypting the message with the receiver's public key which can be only decrypted with the receiver's private key. This traditional method is a form of encryption whereas the digital signatures algorithm is a form of verification. Here the sender encrypts the message with the sender's private key and sends it. The receiver who receives it then utilizes the public key of the sender to decrypt the message. If the message is valid, it implies only the particular sender has sent it thus providing a form of verification.

In RSA the message m is encrypted with the sender's private key d: c ≡ md (mod n) where c is the ciphertext generated. To verify a message, compare the public key e with the result of: m ≡ ce (mod n) where m is decrypted text. This is the RSA Signing and Verification Process.

3.1 Design Diagram

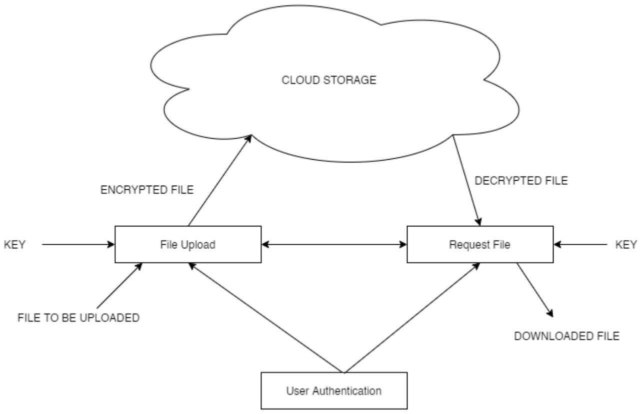

Fig -1: A block-level overview of the model

As represented in Fig. 1, after the user is authenticated with his previously registered email and password, the system provides the user with two options, namely encryption, and decryption.

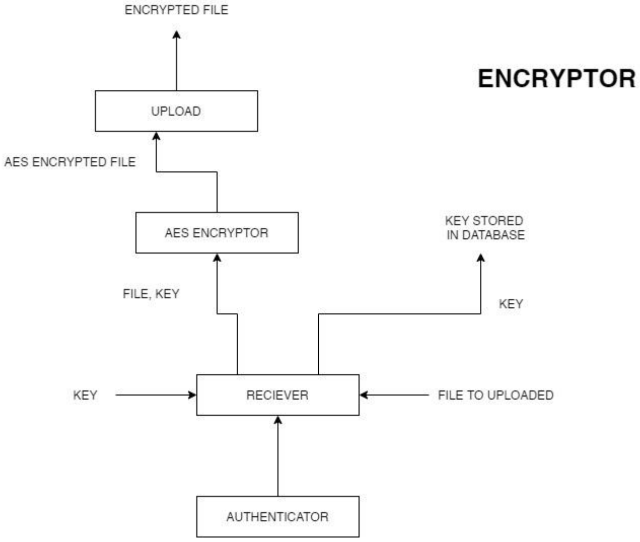

Fig -2: A block-level overview of the encryption process

The inputs for encryption are the file to be encrypted and the secret key. During encryption, the file is split into eight different pieces and encrypted simultaneously with the help of threads. Eight threads simultaneously call the encrypt function which reads the assigned portion of the file and encrypts using the AES algorithm. The key for the AES algorithm is the secret key which is inputted by the user. These fragments are then put together in a zip file and uploaded to the cloud service and the key and file name is uploaded to a highly secure location, either a local database or a secure registry file.

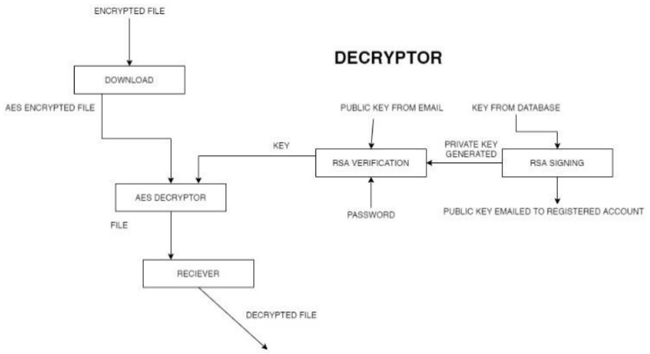

Fig -3: A block-level overview of the decryption process

While decrypting, there is a real-time authentication to avoid brute force attacks. This authentication is in the form of an OTP message using the RSA signing methodology. The filename and the key are extracted from the previously mentioned secure registry file. A public key and private key are randomly generated and the message is signed using the RSA signing algorithm with the private key. The public key is emailed to the registered email. The inputs for decryption are the secret key used during encryption and the public key OTP emailed to the user. While the user inputs both of them, the public key is used for the verification process of the RSA digital signatures algorithm. The message decrypted is verified with the secret key entered and if at any point these yield false, the system prompts an incorrect entry message to the user. If both of these cases are valid then, the user can successfully download the file from the cloud, free from encryption.

4. IMPLEMENTATION, RESULTS AND ANALYSIS

The proposed system was built using Python, HTML and the Flask framework and the cloud service used was Amazon Web Services. An image file was encrypted and successfully uploaded and downloaded from the cloud after providing the necessary keys including the multifactor authentication key or OTP key.

The proposed system stores the zip file on the cloud. In the case of data leakage from the cloud, the fragments from the zip file cannot be restored as it encrypted with AES. Without the key, attempting to decrypt these fragments is futile. Therefore, the security of the system depends on the security of the keys.

The proposed system cannot be attacked via brute force as it requires a verification OTP key which is only emailed to the registered account. Furthermore, the OTP cannot be attacked with brute force as the key length is too big for the time limit considered. Unless the user has access to the system and the email of any registered user and access to the secure database, the chances of an attack are highly unlikely although not impossible.

The proposed system design contains an enhanced AES algorithm. The proposed system performance was compared with the existing AES algorithm's performance concerning space and time complexity and the following results were obtained. Here only the encryption time for both the algorithms are considered.

Table -1: Comparison of Completion Time

File Size (in MBs)	Time is taken by the algorithm (in seconds)	
	AES algorithm	Enhanced AES Algorithm
~1.04	0.009973764419555664	0.006981372833251953
~2.45	0.0179519653320312 5	0.014960527420043945
~3.22	0.022970199584960938	0.020941734313964844
~4.45	0.0284638404846191 4	0.025957345962524414

A line graph visualization of the above table is presented below for better understanding.

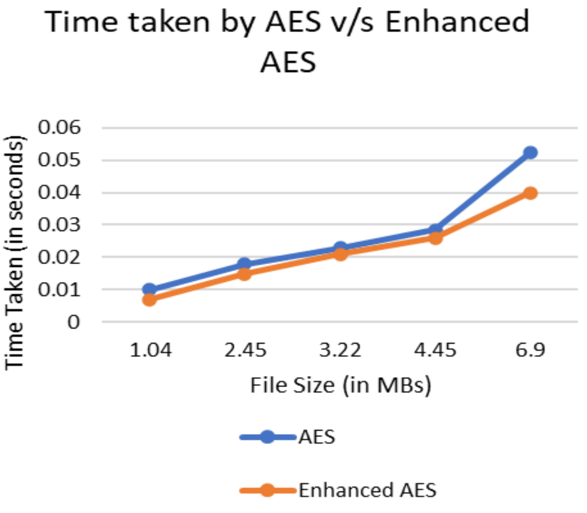

Chart -1: Line graph representation of the comparison of the time taken by both the algorithms

As observed in Chart-1, it is possible to ascertain that the enhanced AES algorithm is much faster than the traditional AES algorithm and thus increasing the overall speed of the hybrid cryptography. As the size increases, the gap between the two curves also increases.

3. CONCLUSION

Protection of sensitive or confidential data is one of the major security concerns in this day and age, for an organization or individual, be it a government body or a business corporation. Cryptography, which initially was designed for military or diplomatic use, can now be publicly utilized by anyone to achieve data security. Encryption is being seen as the best way to make data secure and ensure its protection [6]. Therefore, the utilization of hybrid cryptography can ensure an additional layer of protection to any security system.

The main aim of this paper is to securely store and retrieve files stored on the cloud. Data security is achieved using the combined techniques of AES and RSA Algorithms. Furthermore, the AES algorithm was enhanced with the use of threads. This concept of multithreading reduces the time of completion as the file size increases. It also consists of OTP verification via the RSA digital signatures methodology. Therefore, a system with high security, better performance, data integrity was designed and implemented.

REFERENCES

[1] P. H. Singh, “Design and Implementation of Cloud Services by using Python”, International Journal of Engineering Aad Computer Science, 2018, vol. 07, no. 06, pp. 23971-23981

[2] A. Albugmi, M. O. Alassafi, “Data Security in Cloud Computing”, Fifth International Conference on Future Generation Communication Technologies (FGCT 2016), pp. 55-59

[3] J. Kumar, “Cloud Computing Security Issues and Its Challenges: A Comprehensive Research”, International Journal of Recent Technology and Engineering (IJRTE), 2019, vol. 08, no. 1, pp. 10-14

[4] G. R. Patel, K. Panchal, “Hybrid Encryption Algorithm”, International Journal of Engineering Development and Research, 2014, vol. 02, no. 02, pp. 2064-2070

[5] S. Chaudhari, M. Pahade, S. Bhat, C. Jadhav, T. Sawant, “A Research Paper on New Hybrid Cryptography Algorithm”, International Journal for Research & Development in Technology, 2018, vol. 09, no. 05, pp. 01-04

[6] V. Kapoor, R. Yadav, “A Hybrid Cryptography Technique for Improving Network Security”, International Journal of Computer Applications, 2016, vol. 141, no. 11, pp. 25-30

[7] J. Selvanayagam, A. Singh, J. Micheal, J. Jeswani, “Secure File Storage On Cloud Using Cryptography”, International Research Journal of Engineering and Technology (IRJET), 2018, vol. 05, no. 03, pp. 2044-2077

[8] S. Kanatt, A. Jadhav, P. Talwar, “Review of Secure File Storage on Cloud using Hybrid Cryptography”, International Journal of Engineering Research & Technology (IJERT), 2020, vol. 09, no. 02, pp. 16-20

[9] M. Kumar, S. Chauhan, A. Singh, “An Analysis and Survey of Security Techniques for Cloud Computing”, International Journal of Computer Science and Mobile Computing, 2019, vol. 08, no. 01, pp. 16-20

[10] A. Poduval, A. Doke, H. Nemade, R. Nikam, “Secure File Storage on Cloud using Hybrid Cryptography”, International Journal of Computer Science and Engineering,2019, vol. 07, no. 01, pp. 587-591

[11] B. Bala, L. Kamboj, P. Luthra, “Secure File Storage in Cloud Computing using Hybrid Cryptography Algorithm”, International Journal of Advanced Research in Computer Science, 2018, vol. 09, no. 02, pp. 773-776